Treatment of Data Uncertainties

Nancy M. Larson

Oak Ridge National Laboratory

For analysis of data using either Least Squares or Bayes' Equations (generalized Least Squares), it is necessary to provide appropriate and accurate information regarding the uncertainties on those data. If the data are correlated, then covariances must be provided to the analysis code, and properly incorporated into the fitting procedure. At present, a number of competing methods exist both for generating such covariance matrices and for making use of them. There is no universal consensus as to which methods are correct, nor is there sufficient understanding regarding the consequences of the choices.

In this paper, the generation and use of data covariance matrices will be discussed within the context of the analysis of neutron-induced cross section data via the R-matrix code SAMMY. Two complementary approaches will be described, the first involving mathematical manipulation of Bayes' Equations, and the second utilizing computer simulations.

In the first approach, Bayes' Equations are applied to the raw data (counts per time channel). While it is seldom practical to calculate directly the quantities measured in an actual time-of-flight experiment, nevertheless it is possible to formally express Bayes' Equations in this manner. Further, because raw data are uncorrelated, there is no ambiguity in the treatment of covariances in these equations. The equations are expressed in terms of two distinct types of parameters, those related to the theory (the R-matrix parameters) and those related to the data (the normalizations, backgrounds, and other corrections required in converting from raw to reduced data). The equations are then mathematically transformed (using matrix algebra) and rewritten into different but equivalent equations expressed in terms of reduced data (cross section as a function of energy). The rewritten equations produce results (updated parameter and covariance matrix) exactly equivalent to those which would be produced if one could fit directly to the raw data.

Examination of those equations leads to a definition of the (reduced) data covariance matrix whose use will give results consistent with those given by the analysis of raw data. Essentially, the data covariance matrix must be expressed in terms of the theoretical cross sections, not in terms of the measured cross sections.

Results of the mathematical approach will be confirmed by the second approach: a series of computer simulations will be performed via SAMMY, using a variety of techniques (some of which are correct and others which are not). The simulations will explore the extent to which erroneous results will be obtained by using incorrect/inexact methods.

ORNL is managed by UT-Battelle, LLC for the US Department of Energy under Contract No. DE-AC05-00OR22725.

Email: `LarsonNM@ornl.gov`